

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 22, 2003, 13:48:51 ; Search time 160 Seconds

(without alignments)
5009.446 Million cell updates/sec

Title: US-09-745-506-74

Perfect score: 2789
Sequence: 1 GTGATGTGTTACTTGCTGCT.....TCTGTTACTTACATTCGA 1553

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xl
-Q/cgn2.1/USPTO/US09745506/runat_22082003.132710.10741/apc_query.fasta.1.1735
-DB=SPREMBL_23 -QMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pic -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09745506.ecgn.1.1.223.annat_22082003.132710.10741 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: SPREMBL_23:*
2: SP_archaea:*
3: SP_bacteria:*
4: SP_fungi:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_rv1rus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1373	49.2	340	11	Q8C608 mus musculi

2	707	25.3	159	11	Q9D146	Q9D146 mus musculi
3	358.5	12.9	366	16	Q8R972	Q8R972 thermomane
4	312	11.2	376	16	Q8CSD9	Q8CSD9 staphylococ
5	309.5	11.1	372	16	Q8EP20	Q8EP20 oceanobacti
6	279	10.0	366	16	Q8NMB9	Q8NMB9 staphylococ
7	271.5	9.7	258	16	Q8RE15	Q8RE15 fusobacteri
8	256	9.2	242	5	Q8SV52	Q8SV52 encephalit
9	246.5	8.8	433	16	Q8PKN9	Q8PKN9 corneobacte
10	230	8.2	380	16	Q8NNG6	Q8NNG6 corneobacte
11	203.5	7.3	265	16	Q8DNV4	Q8DNV4 streptococc
12	203.5	7.3	309	16	Q8G5M5	Q8G5M5 bifidobacte
13	189.5	6.8	262	16	Q8DT99	Q8DT99 streptococc
14	185.5	6.7	262	2	P95777	P95777 streptococc
15	181.5	6.5	265	2	Q914G2	Q914G2 lactobacilli
16	174	6.2	262	16	Q8DZAT	Q8DZAT streptococc
17	173	6.2	262	16	Q8E4W8	Q8E4W8 streptococc
18	151.5	5.4	298	17	Q8TMA2	Q8TMA2 methanosarc
19	143.5	5.1	253	16	Q8EUT6	Q8EUT6 mycoplasma
20	138	4.9	306	17	Q8PTJ7	Q8PTJ7 methanosarc
21	135.5	4.9	248	17	Q8TW44	Q8TW44 methanopyru
22	118.5	4.3	1637	6	Q9XSV8	Q9XSV8 bos taurus
23	118.5	4.3	5146	6	Q8SPM4	Q8SPM4 bos taurus
24	114	4.1	250	16	Q8EDX0	Q8EDX0 shewanella
25	114	4.1	1260	5	Q9XCZ5	Q9XCZ5 cryptospori
26	112	4.0	251	2	Q8KTX3	Q8KTX3 vibrio fisc
27	109	3.9	249	16	Q8D2D0	Q8D2D0 wiggleswort
28	106.5	3.8	556	16	Q8RAG1	Q8RAG1 thermomane
29	106	3.8	3352	2	Q9JH83	Q9JH83 streptomyc
30	105.5	3.8	252	17	Q8U1Y9	Q8U1Y9 pyrococcus
31	105.5	3.8	566	16	Q9WZU1	Q9WZU1 thermotoga
32	105	3.8	644	10	Q38736	Q38736 antirrhinum
33	104.5	3.7	456	16	Q8ECAB	Q8ECAB shewanella
34	104.5	3.7	1052	2	Q93CZ8	Q93CZ8 streptomyc
35	103	3.7	655	4	Q92815	Q92815 homo sapien
36	102.5	3.7	623	4	Q75851	Q75851 homo sapien
37	102	3.7	554	16	Q8EJY1	Q8EJY1 shewanella
38	101.5	3.6	850	16	Q8DLR3	Q8DLR3 synechococc
39	101	3.6	238	1	P94953	P94953 methanopyru
40	101	3.6	253	17	Q8RZC0	Q8RZC0 methanopyru
41	101	3.7	488	11	Q8BGE7	Q8BGE7 mus musculi
42	101	3.6	597	5	Q9VVO9	Q9VVO9 drosofila
43	101	3.6	804	16	Q8RF68	Q8RF68 escherichia
44	100	3.6	221	1	Q50523	Q50523 methanobact
45	100	3.6	804	16	Q8XCZ5	Q8XCZ5 escherichia

ALIGNMENTS

RESULT 1

Q8C608 PRELIMINARY; PRT; 340 AA.

AC Q8C608: 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Ngl interacting factor 3-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=1246681;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK076788; BAC36481.1; -
SQ SEQUENCE 340 AA; 37822 MW; 82DC63DB90B6B4E CRC64;

Alignment Scores:

0y 1118 ATGTCGCATCATGACTCTTGAGATCGCTGCTCCCAAGAAATAAATGTCAACCTCTGTGAA 1177
 101 MetSerHisHisAspValLeuAspAlaIleSerIysGlyIleAsnValIleLeuGlySer 120
 0y 1178 CACAGCAACTGAGAGCGCTTCTTCTACCTCGAGATATGCTGAGATTTCACACTG 1237
 121 HisSerAsnThrGluArgGlyPheLeuSerIleuGlnGluMetLeuGlyValHisPhe 140
 0y 1238 GAGATATAAGATAATATATATCTTATCAGAGACTGACAGAGACCTCTTCAGAGTGTA 1294
 141 GluAsnIysIleAsnIleIleLeuSerGluThrIAspArgAspProLeuAlaGlyVal 159

RESULT 3

ID	Q8R972.	PRELIMINARY;	PTM;	376 AA.
AC	Q8R972.			
DT	01-JUN-2002 (TReMBLrel. 21, Created)			
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)			
DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)			
DE	Hypothetical protein TTEI1753.			
GN	TTEI1753.			
OS	Thermoanaerobacter tengcongensis.			
OC	Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;			
OC	Thermoanaerobacteriaceae; Thermoanaerobacter.			
OX	NCBI_TaxID=119072;			
RM	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-M84 / JCM 11007;			
RX	MEDLINE=21992816; PubMed=11997336;			
RA	Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,			
RA	Tan H., Chen K., Wang J., Yu J., Yang H.;			
RA	Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,			
RT	"A complete sequence of T. tengcongensis genome.";			
RL	Genome Res. 12:689-700(2002).			
DR	EMBL; AE013129; AAM24947.1; -			
DR	InterPro; IPR002678; DUF34.			
DR	Pfam; PF01784; D0F34; 1.			
DR	TIGRFAMS; TIGR00486; TIGR00486; 1.			
KO	Hypothetical protein; Complete proteome.			
SO	SEQUENCE 376 AA; 41674 MW; 557753F2A1E745C6 CRC64;			

Alignment Scores:

Pred. No.:	1.91e-23	Length:	376
Score:	358.50	Matches:	105
Percent Similarity:	48.47%	Conservative:	85
Best Local Similarity:	26.79%	Mismatches:	123
Query Match:	12.85%	Indels:	79
DB:	16	Gaps:	14

US-09-745-506-74 (1-1553) x Q8R972 (1-376)

```

QY 254 AAGCGTCCGCTCTCTCTCTCTGAAAGACATCTCCGCTGGTTTGGAGATGGAGC 313
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 GlnThrIleValSerIleMetAspArgIleuAlaProAlaGlyuSPheAlaGlnAspTyrAsp 30
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 AATGTGGATTAAGTGGTGGAAACAAGCCCAACATACCTGTAATATACACTCTTCCGACC 373
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 AsnValGlyIleuIleVal---GlyAspProIeuGlnAspValSerIlyIleuValAla 49
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 374 AATGACCGATGAGAAAGATGATGGAGAGGCTGCAAAAAGAGCGACAGCTATCTGC 433
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 50 LeuAspAlaThrIleGluValAlaGlnIleAlaIleGluIlyuIysValAspMetIleVal 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 434 TCGTACCACTCCGCGCTATCTTCCGACCCATACGCGCATACCTGGAAACAATCGAAGAG 493
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 ThrIleAspProIeuIlePheIleuSerIleuIysSerValAlaGlySerAspThrProValGly 89
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 494 CGCGCTGGATCCGAGCGCTGCGAACAAGATCGATATCACTCTCCATACAGCGCTAT 553
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 90 MetIeuValMetIyIleuIleIeuAlleGluIleSerIleuGlyAlaIleAlaIleThrSerPhe 109
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 554 GATGCTCGGCGCCCAAGGCGTCAACAACAGCTGGTGGCTAAAGGCTTGGAGCTTGAATCTCC 613
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Dd	110	AspIIlealPAr6a	nglGlymetAsnApRIleucysanlleu	gylle-----	126
Oy	614	AGGCCATAC	TCCTCCAAAGCTCCCAACTACCTCAGAGGGAACCA	CCGAGTAGAA	673
Dd	127	-----	-----	-TyrGluArgGluVal	131
Oy	674	TTCAACGTAACT	CACCCCAAGACCTGGACAAATCTGTCGACAGC	---AAAGCAATT	730
Dd	132	LeuAspValThr	TyrSerGluAspTyrGlySerValValValValTyrValPro	lysGlyTyr	151
Oy	731	GACGGTGT	-----	---TCT	742
Dd	152	GluAspValVal	AlaArgAsnAlaMetCysSerAlaGlyAlaGlyPheIle	GlyAsnTyrSer	171
Oy	743	GTCACCTCTTT	CTGCTAGC---ACTGCT	-----	769
Dd	172	AspSerThrPhe	GluValGluGlyThrGlyThrPheLysPromet	GluGlyAlaAspPro	191
Oy	770	-----	-----AAAGGAGCAAAACCGGATTAATCGAAATTG	802	
Dd	192	PheIleGlyGluVal	GlyLysLeuGluLysValGluGluIleAlaGluThrValVal	211	
Oy	803	ACTCAGAGCGCT	TGATGAGCGGTGATTTCTTCCGCG-----	AAACAA	850
Dd	212	ProGluInsTyr	LeuAspLysValIleAsnAlaMetLeuArgValHisProTyr	GluGlu	231
Oy	851	CAACTTATATCA	GAAGAGGAATTTCTGATCTGACAGAGACCTTGTCTTACATATCGA	910	
Dd	232	ValAlaTyrAsp	ValTyrSerLeuAlaAsnLeuArgLys-----	GluTyrGly	247
Oy	911	ATGGAGCGGTT	ATGCACATCGAATCTGTCTCCCTGGCAACCATGATGATGCAATA	970	
Dd	248	LeuGlyArgIle	GlyValIleGlyGluThr---ThrLeuLysGluLeuAlaLeuLysVal	266	
Oy	971	AAAGAGACAC	CTAAATATCTCATATTCGCTTACCCCTGGCGGTGGGAGAACCTTAGAG	1030	
Dd	267	LysAlaLysLeu	LysValAlaGluSerLeuLysVal-----	ValGly--AspLeuGlu	282
Oy	1031	TCTCAAGTCAA	AGTCGTGGCCCTGTGCTGCTGTTCTGGAGCAC-----	1075	
Dd	283	ArgArgValVal	LysLysValAlaValLysGlyLysGlyAlaSerLeuIleHisLysAla	302	
Oy	1076	GTTCTGCAGGG	TGTGGAGCTGACCTTATCTCAAGGTAGAGATGCTCCATCATGATCT	1135	
Dd	303	ValMetLysGly	-----AlaAspValLeuIleThrAlaAspIleGlyTyrHisAspAla	320	
Oy	1136	TTGGATGGCTG	CTCCAGAAATAAAGTCACTCTGTGTACACAGACACAGACAGAGCA	1195	
Dd	321	ValGluAlaGln	HisLeuGlyLeuAlaLeuIleAspAlaGlyHisPheAlaThrGluAsn	340	
Oy	1196	GCGTTTCTT	CTGACCTTCGATATGCTGGATTCTCATTGGAG-----	AAATAG	1246
Dd	341	IleAlaValAla	ArgPheIleAlaGluTyrLeuLeuAspGluThrGlnLysGlnGlyHisGlu	360	
Oy	1247	ATTAATATTA	CTATACAGATGACATGACGAGGACCT	1282	
Dd	361	IleGluValAla	PheValSerGluValGlnLysAspPro	372	
RESULT 4					
QCSD9					
AC	QCSD9	PRELIMINARY;	PRT;	366	AA.
DT	01-MAR-2003	(TREMBLrel. 23, Created)			
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE	Conserved hypothetical protein.				
GN	SEI246.				
OS	Staphylococcus epidermidis.				
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
OX	NCBI_TaxID=1282;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 12228;				

RESULT 4

ID	08CSD9	PRELIMINARY;	PRT;	366 AA.
AC	08CSD9;			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Conserved hypothetical protein.			
GN	SEI246.			
OS	Staphylococcus epidermidis.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1282;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 12228;			

RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
 RA Chen Z., Wen Y.,
 RA Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AE016748; AA004845.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 366 AA; 41859 MW; BF3A88EE77462A1 CRC64;

Alignment Scores:

Pred. No.: 3,02e-19 Length: 366
 Score: 312.00 Matches: 100
 Percent Similarity: 44.78% Conservative: 80
 Best Local Similarity: 24.88% Mismatches: 130
 Query Match: 11.19% Indels: 92
 DB: Gaps: 16

US-09-745-506-74 (1-1553) x Q8CSD9 (1-366)

```

OY 245 ATGATTTGAAAGGCTCTCTCTCTCTCTGTAATGACTTTGCATTCCTCTCTGTTGCTGAG 304
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 1 MettysileserGluLeuMetGluValLeuAsnAsnHisValProPheHisGlnAlaGlu 20
OY 305 AGTTGGACAAATGTGATTTGATTTGCTGAGAACCAACCCACATACCTGTAATACACGC 364
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 21 SerTrpAspAsnValGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 39
OY 365 TTCTGACCAATGACCTGACTGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 40 LeuThrThrLeuAspCysThrAspAspValValAsnGlnAlaIleGluLeuAsnThrAsn 59
OY 425 CTCATCTCTCTCTACATCCGCTATCTTCGACCAACCAAGAGAGAGAGAGAGAGAGAGAGAG 484
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 60 ThrIleIleAlaHisHisProLeuIlePheLeuGlyValValValValValValValValVal 79
OY 485 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 80 TyrGlySer---IleIleArgGlySerLeuIleGlnAsnAsnIleAsnLeuIleAlaLeuHis 98
OY 545 ACAGCCTTGATGCTGCGCCGCCAGGCGCTCAACCAATGCTGTTGAGGCTTGGAGCT 604
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 99 ThrAsnLeuAspValAsnProLysGlyValAsnArgMetLeuAlaAspGlnIleGlyLeu 118
OY 605 TGTACTCTCCAGGCCATCATCTCTCCCAAGCTCCCACTACCTACCTACCA 652
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 119 GluAsnIleSerMetIleAsnThrAsnSerSerTyrTyrTyrLysValGlnThrPheIle 138
OY 653 -----GAG 655
OY 139 ProLysAsnTyrIleGluAspPheLysAspSerLeuAsnGluLeuGlyLeuAlaLysGlu 158
OY 656 GGAACCAACCGAGTAGAATTCAACGTTAAC----- 685
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 159 GlyAsnTyrGluTyrCysPhePheGluSerGluGlyLysGlyGlnPheLysProValGly 178
OY 686 -----TACACCCAGACCTGACGACCAAGTCATCTGCACTGGAAGAAATGAC 733
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 179 AspAlaSerProTyrIleGlyLysLeuAsp-----SerIleLeuTyrValAsp 194
OY 734 GGTGTTCTGTCATCTTTCTTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 793
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 195 GluIleLysLeu---GluPheMetIleLys---GlyAsnGlu----- 206
OY 794 CTGAATTTGACTCAGAGAGCTTTGATGCAG-----GTGCTAGAT 832
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 207 LeuGluIleThrLysArgAlaIleLeuAspAsnHisProTyrGluThrProValPheAsp 226
OY 833 TTTCTTCCCGGAGAACAACTTTATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 227 PheIleLysMetAsnLysGlu----- 233
OY 893 TTGCTTCTACATACGAGATGAGAGAGAGATGACACATGAGATGATCTCTCCCTGCA 952
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 234 -----SerGluTyrGlyLeuGlyIleIleGlyGlnLeuAsnGlnThrThrLeuAsp 251

```

```

OY 953 ACCATGATTCATCGAATAAAGACACCTAAACTATCTCATATTCGTTAGCCCTTGG 1012
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 252 GluPheSerGluTyrAlaLysGlnLeuAsnIleProSerValArgTyr----- 268
OY 1013 GTGGGAGAACCTTGTAGCTTCATCAAGTCGAGGCGCTGTGCTGCTTGGAGAGC 1072
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 269 ThrGlyGlnHis---AspSerProIleLysValAlaIleIleGlyLysSerGlyLe 287
OY 1073 AGCGTTTCGACAGGCTGTGAG-----GTCGACCTTACCTCA 1111
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 288 -----GlyPheGluTyrTyrLysAlaSerGlnLeuGlyAlaAspValPheValThr 303
OY 1112 GGTGAGATTCCTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 304 GlyAspIleLysHisHisAspAlaLeuAspAlaLysIleGlnAsnValAsnLeuAsp 323
OY 1172 TGTGACACGACCAACACTGACGAGAGCTTTCTTTGACCTTCGAGATATGCTGATTC 1231
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 324 IleAsnHis---TyrSerGluTyrValMetLysGluGlyLeuLysGluLeuLeuLys 342
OY 1232 CACTG-----GAGATTAAGATTAATATTCATGACAGACTGACAGGAGCCCT 1282
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 343 TrpLeuPheLysTyrGluAsnGlnPheProIleTyrAlaSerGluIleAsnThrAspPro 362
OY 1283 CTTCAG 1288
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 363 PheLys 364

```

RESULT 5

Q8EP20 PRELIMINARY; PRT; 372 AA.

AC Q8EP20;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE Hypothetical conserved protein.
 GN OB1940.
 OS Oceanobacillus thelyensis.
 OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE=22220767; PubMed=12235376;
 RA Takami H., Takaki Y., Uchiyama I.,
 RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments.",
 RL Nucleic Acids Res. 30:3927-3935(2002).
 DR EMBL: AF004599; BAC13896.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 372 AA; 41869 MW; B66B9C9C8749C43 CRC64;

Alignment Scores:

Pred. No.: 5.11e-19 Length: 372
 Score: 309.50 Matches: 105
 Percent Similarity: 46.79% Conservative: 70
 Best Local Similarity: 28.07% Mismatches: 152
 Query Match: 11.10% Indels: 47
 DB: Gaps: 13

US-09-745-506-74 (1-1553) x Q8EP20 (1-372)

```

OY 269 TCCCTGAATGACTTGTGACCTCTCTGCTGAGAGTTGGAGCAATGTTGATTAAGT 328
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 14 AlaMetAspValTyrPheLysHisLeuAlaTyrAspTrpAspAsnValGlyLeuGln 33
OY 329 GTGGAACCAAGCCACCACTGTAATACACTCTTCTGACCAATGACTGACTGAG 388
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 34 IleGlySerPheAsnAlaIleThrSerLysValLeu---IleThrLeuAspValThrGln 52
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
OY 389 GAAAGTATGAGAGAGAGTGTGCAAAAGAGGAGACCTCATCTCTCTACATCCGCT 448
DB      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

Db	53	SeValValaSpArgLualaIleIleGluIbysGlyMetAsnLeuIleIleAlaH::SHisSerLeu	72
OY	449	ATCTCCGACCCATGAAGCGCATTAACCTGGCAACACATGGAAAGACGCCCTGGTATCCGG	508
Db	73	LeuHeGlnProIleuIlySGlnIleAsnMetAspThrProIySGlyArgThValGlnIlys	92
OY	509	GCTGTGGAGAACAGAGCGGTATCTACTCTGCTATACAGGCTTAATGATGCTGGCGCCGAC	568
Db	93	LeuIleIysAsnAspIleThrValTYrAlaSerHisThrAsnLeuAspAlaAsnAsn	112
OY	569	GGCGTCAACAACATGGTGGCTTAAGGCGCTTGGACCTTGTACTCCAGGCCCATACATCT	628
Db	113	GlyValAsnAspMetLeuAlaThrGluLeuIlysLeuThrHisThrIleProIeu-----	130
OY	629	TCCAAAGCTCCCAACTACCTACAGAGGGAAACCAACGAGTAAATTCACAGTTAACTAC	688
Db	131	-----ValProValTYr--AsnGlnIlysAsnTYrIlyVal-----ValIleTYr	144
OY	689	ACCCAGAC-----CTGACAAAGATCTATGCTGACAGTAAAGAAATTCAGCGTCTTCT	742
Db	145	ValProGluSerHisIleGluIlySIleArgGlnAlaPheAlaGluSerGlyAlaGlyTYr	164
OY	743	GTCACCTCTTTTCTGTACGACT-----	766
Db	165	IleGlyAsnTYrSerHisCysThrPheGlnSerProGlyGlnGlyThrPheIysProIeu	184
OY	767	-----GGRNATAG-----	787
Db	185	GluGlyThrAspProTYrIleGlyGluGlnAspIlySIleThrTYrValAspIuIyIlys	204
OY	788	ATTATATCGAATGTGTACTACAGAGGCTTTGTATGACAGTGA--GATTTTCTTTCCCGG	844
Db	205	IleGluThrIleValProGluSerIleLeuSerAsnValIyISerIleLeuSerAla	224
OY	845	AACAAACACATTTCAGAAACGGAATTCGTGCACTGGAGAAACCTTTCCTTACAT	904
Db	225	HisProTYrGluGluProAlaTYrAspIleTYrProLeuGluAsnIly--ThrThrAsn	243
OY	905	ACTGAATGGAGCGGTATGCAACACTGATGAATGTCGCCCTGGCAACATGATTTGAT	964
Db	244	PheGlyIleGlyAlaGlyIleGlyAsnLeuThrGluProThrIleAspSerPheIleIys	263
OY	965	CGAATTAATAAAGACACTAAACTATATCTATATTCGCTTACGCCCTGGGGTGGGAGAAC	1024
Db	264	ThrValIySIleThrGlnPheGlnLeuGlnGlyIleArg-----IleSerGlyAsn	279
OY	1025	TTTAAGTCTACAGTCAAAAGTCGTGGCCCTGTGTCTGTGGTCTGGAGACACGTTCTGCAG	1084
Db	280	LysAspIlySIlyIleIyAspValAlaIleLeuGlnGlyIySerGlyGlyIyValIleSer	299
OY	1085	GGGTGT-----GAGCGTACCTTTACTCTACACAGGTGAGATGTCCCAATCATATACT	1133
Db	300	HisAlaMetArgIyLysIyAlaAspAllyIleThrIyIyAspMetThrPheHisIleAla	319
OY	1136	TTTGGATGCTGCTCCCAAGAAATTAATGTCACTCCCTGTGAACACAGACACACTAGAGA	1199
Db	320	GlnAspAlaAlaGluMetGlyLeuThrAlaIleAspAlaGlyHis---TYrIleIuIys	338
OY	1196	GGCTTTCTTTGTGACCTTCGAGATATGCTGGATTTCTCACTTGGAC--ATATAGATAAAT	1255
Db	339	ValMetIySIyTYrTYrThrIyGlnIyIyLeuAlaGlnThrMetGluLeuAspIyAspPhe	358
OY	1253	ATTATCTATACAGACACTGACAGGAGCGCTCTGACGGGTA	1294
Db	359	IleGluValISerGluIleAsnThrAspProPheGlnPheVal	372

RESULT 6		
O8NMWB9		
ID	O8NMWB9	PRELIMINARY;
AC	O8NMWB9	PRT; 366 AA.
DT	01-OCT-2002	(TRENABLE). 22, Created)
DT	01-OCT-2002	(TRENABLE). 22, Last sequence update)
DT	01-MAR-2003	(TRENABLE). 23, Last annotation update)

DE Hypothetical protein MM1511.
GN MM1511.
OS *Staphylococcus aureus* (strain MR2).
OC Bacteria; Firmicutes; Bacillales; *Staphylococcus*.
OX NCBI_Taxid=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda H., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA".
RL Lancet 359:1819-1827(2002).
EMBL: AP004827; BAB95376.1; -.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; DUF34. 1.
DR TRIFRAMS: TRIGR00486; TRIGR00486; 1.
KW Hypothetical protein, Complete proteome.
SQ SEQUENCE 366 AA; 41153 MW; FA1E287620876C79 CRC64;

Alignment Scores:	
Pred. No.:	2,9e-16
Score:	279.00
Percent Similarity:	43.40%
Best Local Similarity:	25.89%
Query Match:	10.00%
DB:	16
Length:	366
Matches:	102
Conservative:	69
Mismatches:	151
Indels:	72
Gaps:	16

[illegible]

US-09-745-506-74 (1-1553) x Q8FNK9 (1-433)

OY 194 ACAGACGCGG-----TTGTAGATTCCTGATCTGCAATTTCTCCGTTCTTCAG 247
 |||||:||||| ||| ||| |||||: |||
 Db 47 ThlrhmetAalgglyProhIsValSerAspLeuThrIleGlyHisIlaArgThrIleMet 66
 OY 248 GATTGAAGGCTCTCTTCTCTCTCTGTAAGTACCTTGACCTCTGCTGAGAGT 307
 |||||:||||| ||| ||| |||||: |||
 Db 67 -----AsnGlnAlaTyProProAlaLeuAlaGlnSer 77
 OY 308 TGGGACATGTTGGATTACTGTG---GACCAAGCCCAACACACTGTAATCACTC 364
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 78 TrpAspLysValIleGlyLeuIleCysGlyAspProGlyLysPro-----ValArgArgVal 95
 OY 365 TTCTGACCAATGACCTGACTAGAGAGTGTGAGAGAGTGTGCAAAAGAGGACACAC 424
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 96 AlaLeuAlaLeuAspCysThrGlnGluValAlaAspArgAlaValGlnLeuGlyValAsp 115
 OY 425 CTCATTCCTCTCTACCATCCGCTATCTTCGAGCCACGAGGACATACCTGAGACACA 484
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 116 MetLeuIleIleHisIleAspProLeuLeuArgGlyValIleSerValAlaAlaAspGlu 135
 OY 485 TGAAGAGAGCGCTGTGATCCGGCTCTGAGAACAGAGTGTGATCTGCTGCTAT 544
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 136 ProLysGlyArgValIleHisThrLeuIleArgGlyGlyValAlaLeuPheSerIleHis 155
 OY 545 ACAGCCTATGATGCTGCGCCCGAGGCGTCAACAACCTGTTGGCTAAAGGCTTGAGCT 604
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 156 ThrAsnAlaAspSerAlaArgProGlyValAsnAspLysLeuAlaGlnLeuValIle 175
 OY 605 TGTACTCTGAGCCCATCATCTCTTCCAAAGCTCCCAACTACCTACAGAGGAAACAC 664
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 176 ThrProGlyArgProIleAlaProArgTyLeuAspAlaLeuAspLysTrpGlyValHis 195
 OY 665 CGAGTAGAA----- 673
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 196 ValLeuGlyLysAspAlaArgGlnValLysArgAlaLeuPheAspAlaGlyAlaGlyLys 215
 OY 674 -----TTCAACGTTAACTACACCCCAAGACCTGACCAAGCTC 709
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 216 IleGlyGlyTyThrGlnGlyValPheGlnLeuAsnGlyThrGlnPheArgProVal 235
 OY 710 ATGCTGCA-----GTGAAGGAATTGACGCTGTTCTGCTACTCTTTTCTGCTAGG 763
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 236 GlyLysAlaAspProThrGlnGlyThrValGlyLeuPheArgAspSerGlnLeuArg 255
 OY 764 ACTGATATGAGGAACAACAGGATTAATCGAATGTACTGACAGAGCTTGATGAG 823
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 256 ValGlnPheValAlaProArgArgLeu-----ArgAlaArgLeuGln 270
 OY 824 GTGTAGATTTCTTCCCGAACAAACACTTATTCAGAGAGC-----GAAATTCG 877
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 271 ValVal-----ArgGlnAlaHisProTyGlnGlnProAlaPheAspValVal 286
 OY 878 TCACGTGGAAGACCTTCTCTTCTACATCT-----GGAATGGACGCTTATGACACTG 931
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 287 GlnLeuProAlaThrAspAlaLeuGlnTyThrAlaTyGlyLeuGlyArgValGlyGlnLeu 306
 OY 932 GATGAATCTGTCTCCGCAACATGATGTGATGCAATAAAGACACTTAAACTATCT 991
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 307 ProGlnProMetArgLeuAlaAspPheValGlnGlnValAlaAspGlnLeuProAlaThr 326
 OY 992 CATATTGCTTACGCTTGGGTGGGAGAAC-----TTAGAGTCAATCAATCAATGTC 1045
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 327 -----ValTrpGlyValAlaArgAlaThrGlyAspProAspGlnMetValSerThr 342
 OY 1046 GTGGCCCTGTGTGTGTGTGTGGAGACGCTTGTGAGGGTGTGAG-----GCT 1096
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 343 ValAlaValSerSerGlySerGlyAspSerPheLeuAspArgValArgGlnLeuGlyVal 362
 OY 1097 GACCTTACCTCAACAGGAGATGCTCCCATGATGATCTTGGAT-----GCTGCTTCCCA 1153
 |||||:|||||:|||||: ||| ||| |||||: |||

Db 363 AspValTyValThrSerAspLeuArgHisIleProValAspGlyTyLeuAlaGly 382
 OY 1154 GGAATAATGTCACTCTGTGACACAGCAACTGACAGAGGCTTCTTCTGACCTT 1213
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 383 GlyProAlaValAlaAspThrAlaHisIleThrPheAspGlnPheProThrAlaGlnVal 402
 OY 1214 CGAGTATGCTGGATTCCTACTTG---GAGAAATAAGTAAATATATCTTACAGAGCT 1270
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 403 ArgAspLeuLeuAlaAlaGlnAlaProGlyValSerValGlyValLeuAsnIleArgThr 422
 OY 1271 GAC 1273
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 423 Asp 423

RESULT 10
 Q8NNG6 PRELIMINARY; PRT; 380 AA.
 ID Q8NNG6;
 AC Q8NNG6;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DF 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Uncharacterized ACR.
 GN CGL2238.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxId=1718;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005281; BAB99631.1;
 DR InterPro; IPR002678; DUF34.
 DR Pfam; PF01784; DUF34.1.
 DR TIGRFAMS; TIGR00486; TIGR00486.1.
 DR Complete proteome.
 SQ SEQUENCE 380 AA; 41155 MW; DF995E1E33E84516 CRC64;

Alignment Scores:
 Pred. No.: 7,886-12 Length: 380
 Score: 230.00 Matches: 89
 Percent Similarity: 44.35% Conservative: 72
 Best Local Similarity: 24.52% Mismatches: 154
 Query Match: 8.25% Indels: 48
 DB: 16 Gaps: 13

US-09-745-506-74 (1-1553) x Q8NNG6 (1-380)

OY 293 TCGTTTGTGAGAGTTGGGACAAATGTGATTAATCTGTGGAACCAAGCCCAACATCT 352
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 20 AlaLeuAlaGlnSerThrAspLysValGlyLeuIleCys---GlyAspProThrGlnSer 38
 OY 353 GTAATATACCTTCTCTGACCAATGACCTGAGAGTGAAGAGTGAAGAGAGTGTGCA 412
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 39 ValLysArgValGlyLeuAlaLeuAspCysThrGlnAlaValAlaAspLysAlaValAsp 58
 OY 413 AAGAAGGACAGCATATCTCTCTACCATCGGCTATCTTCCGACCAAGAGGACATA 472
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 59 MetGlyLeuAspMetLeuIleIleHisIleProLeuLeuAlaGlyValIleThrSerAl 78
 OY 473 ACSTGAAACATGGAAGAGAGCGCTGTGATCCGGCTGTGGAACAGAGTGGTATC 532
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 79 AlaAlaAspGlnProGlyGlyValIleValIleHisIleThrLeuIleArgGlyValAlaLeu 98
 OY 533 TACTCTCTCAADACAGCTTATGATGCTGCGCCCAAGGCGGCAACACTGTTGGCTAA 592
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 99 PheSerAlaHisIleThrAsnAlaAspSerAlaArgProGlyValAsnAspLysLeuAlaGlu 118
 OY 593 GGGCTTGAGCTTGTACTCTGACGCGGCATATGCTTCCCAAGTCCCAATACCGCTCA 652
 |||||:|||||:|||||: ||| ||| |||||: |||
 Db 119 LeuValGlyIleThrAlaGlyArgProIleAlaThrArgLeuLeuGlyIleMetAspLys 138


```
QY 1040 AAGTCGCGCCCTGTGCTGCTTCTTGAGACAGCCTTTCGAGGCTGT----- 1090
DB 176 SerAlaValAlaIleCysGlySerGlnSerPheTyrLysAspAlaLeuAlaLys 195
QY 1091 GAGCGTCACCTTTCACCTCAGAGTGAATGCCATGATCTTGTGATGCTGCTTC 1150
DB 196 GAlaAlaAspValTyrIleThrGlyAspIleTyrThrAlaGlnAspMetLeuSer 215
QY 1151 CAGGAAATAAATGTCATCTCTGTGACACAGCAACTGACAGGCTTCTTCTGAC 1210
DB 216 AspGlyLeuLeuAlaLeuAspProGlyHis--TyrIleGluValLeuPheValGluLys 234
QY 1211 CTTCGAGATATGCTGATCTTCACCTGAGAAATAG-----ATTAATATTTACCTA 1261
DB 235 IleAlaAlaLeuLeuThrGlnTyrPlysgLysGlyTyrGluLeuGluIleuPro 254
QY 1262 TCAGAGACTGACAGGACCTCTTCAG 1288
DB 255 SerGlnAlaSerThrAsnProPheArg 263
```

RESULT 12

```
Q8G5M5 PRELIMINARY; PRT; 309 AA.
ID O8G5M5
AC O8G5M5
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein with duf34.
GN BL0984.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NC 2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Schell M.A., Karimiantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwiolen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL, AE014721; AAN24792.1;-.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 309 AA; 33851 MM; 683448905FP0F3 CRC64;
```

Alignment Scores:

```
Pred. No.: 1,81e-09 Length: 309
Score: 203.50 Matches: 67
Percent Similarity: 36.48% Conservative: 45
Best Local Similarity: 21.82% Mismatches: 106
Query Match: 7.30% Indels: 89
DB: 16 Gaps: 5
```

US-09-745-506-74 (1-1553) x Q8G5M5 (1-309)

```
QY 248 GATTGAAGCTCTCTCTCTCTCTGTAATGACTTTGCATCCCTCTGTCGAGAGT 307
DB 6 AsnLeuLysGlnValIleAspValLeuGlnThrLeuTyrProLeuArgTyrAlaGluGln 25
QY 308 TGGGCACTGTTGATCTAGTGTGAGAACCAAGCCACACATACGTAAATACATCTTC 367
DB 26 TTPASPGIUPROGlyLeuIleVal--GlyAspLeuSerHisAspValHisArgIleVal 44
QY 368 CTGACCAATGACCTGACGAGAGTGAAGTGAAGAGTGTGCAGAAAGAGCAGACCTC 427
DB 45 PheAlaAlaAspProThrSerAlaIleIleAspLysAlaIleAlaThrGlyAlaAspLeu 64
QY 428 ATTCTCTCTACATCCGCTATCTTCGACCCATGAAGCCCATACCTGGAACATG 487
DB 65 LeuIleThrHisHisProLeuPhePheArgSerValHisGluThrSerGlyLeuGlyPhe 84
```

```
QY 488 AAGAGCCCTGTGATCCGGGCTCTGAGAAACAGAGTCGTATCTACTCTCATCA 547
DB 85 ArgGlyAspIleValAlaArgAlaLeuTyrGlnHisGlyCysGlyLeuTyrValGlyHisThr 104
QY 548 GCGTATGATGTGGCGCCCGGCGGTCAACAACCTGTTGGCTTAAGGCTTGAGACTGT 607
DB 105 AsnAlaAspAlaIleTyrArgGlyValGlyAlaIleAlaIleAspTyrPheGlyLeuIle 124
QY 608 ACCTCCAGCCCATCATCTCTCCCAAGCTCCCACTACCTACAGAGGAAACCCAGA 667
DB 125 AspGlnLysProLeuValProIleAspAspAlaAsn-----AlaSerHisPro 140
QY 668 GTACATTTCAACGTTAACTACACCCAGACCTGGCAAGATCATGTCTGCAAGTAAAGA 727
DB 141 Val----- 141
QY 728 ATTGACGGTGTTCGTACATCTTTTCTGTAGACTGGTAAATAGAGAACAAACACGG 787
DB 141 ----- 141
QY 788 ATTAATCTGAATTTACTACAGAGGCTTTGATGAGTGTAGATTTTCTTCCGGAAC 847
DB 141 ----- 141
QY 848 AAACACTTTATCAGAACAGGAAATTTCTGCATCGAGAAAGCTTTGCTTACATACT 907
DB 141 ----- 141
QY 908 GGAATGGAGCGGTTATGACACATGTGATCTGCTCCCGCAACCATGATGATGCA 967
DB 142 GlyLeuGlyArgValGlyArgLeuProLysProValAlaLeuLysAspPheAlaGlnArg 161
QY 968 ATAAAGACACCTAAACCTATCTCATTTGCTTACGCTTGGGGGTGGAGAACCTTA 1027
DB 162 ValPheAspGluValSerAspHisGlyMetThrSerAlaLeuGlyIleGlnValCysGly 181
QY 1028 GAGCTCTCA-----GTCAAGTCTGGCCCTGTGCTGCTTGGAGACAGCTTCTG 1081
DB 182 AspSerAspThrPheValGlnTyrValAlaIleLeuProGlySerGlyAspSerLeuPhe 201
QY 1082 CAGGCTGTGAGGCT-----GACCTTACCTCAGACGAGATGTCCATCATGAT 1132
DB 202 AsnGluValAlaArgAlaThrGlyValAspValIlyrValThrSerAspLeuArgHisHisPro 221
QY 1133 ACTTTGATGCTGCTTCCCA 1153
DB 222 ValThrAspAlaIleGluGln 228
```

RESULT 13

```
Q8DT99 PRELIMINARY; PRT; 262 AA.
ID Q8DT99
AC Q8DT99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN SMU.1463C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=DA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans DA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL, AE014978; AAN59121.1;-.
KW Hypothetical protein; Complete proteome.
```

SQ SEQUENCE 262 AA; 29740 MW; A5740C55B8EF248B CRC64;
 Alignment Scores:
 Pred. No.: 3,15e-08 Length: 262
 Score: 189.50 Matches: 66
 Percent Similarity: 36.59% Conservative: 50
 Best Local Similarity: 20.82% Mismatches: 100
 Query Match: 6.79% Indels: 101
 DB: 16 Gaps: 10
 US-09-745-506-74 (1-1553) x Q8D199 (1-262)
 OY 353 GTAATAACACTCTTCCTGACACATGACCTGAGAGAGTGTGAGAGAGAGTGTGCA 412
 Db 36 lIeLysArgLeuMetIleAlaLeuAspValArgIuThrThValAlaIleGlu 55
 OY 413 AAGAAGGACAGCACTTCTCTCTACACATCCGCTATCTCCGACCACTGAAGCGCAT 472
 Db 56 LysLysValAspLeuLeuIleValLysHisAlaProIlePheArgProLeuLysAsnLeu 75
 OY 473 ACCTGGAACACATGGAAGAGGCGCTGATCCGGCTGGAGAACAGAGTGGTATC 532
 Db 76 Val---GluThrAlaGlnAsnHisIleTyPheAsnLeuIleLysHisAlaVal 94
 OY 533 TACTCTCTCATACAGCCTATGATGCTGGCCCGCCAGGCGCTCAACACAGTGTGGCTAA 592
 Db 95 TyValSerHisThrAsnIleAspIleValProAspGlyLeuAsnSprTrpPheCysAsp 114
 OY 593 GGGCTTGGAGCTTGTACCTCCAGGCCCATACATCTCTCCAAAGCTCCCAACTACCTTACA 652
 Db 115 LeuLeuAspIleLysAsnArgIleLeuSerProSerLysAsp----- 129
 OY 653 GAGGAAACACAGAGTGAATTCACAGTAACTACACCAAGACCTGGACAAAGTCATG 712
 Db 129 ----- 129
 OY 713 TCTGACGTAAAGAAATGACGGTGTCTGTCACTTTCTTCTGTAGAGACTGTAA 772
 Db 129 ----- 129
 OY 773 GAGGAACAAACAGGATTATCTGATTTGACTCAGAAAGCTTGTGATGAGTGTAGAT 832
 Db 130 ----- 130
 OY 833 TTTCTTCCCGGAACAACACTTATCAGAAAGCAAAATTCGTCACTGAGAGAGCGT 892
 Db 131 Tyr----- 131
 OY 893 TTGCTTCTACATACGTGAATGGAGCGTATATGACACATGATGATCTGTCCCTGGCA 952
 Db 132 ----- 132
 OY 953 ACCATGATGATGCAATTAAGAAACACTTAACATTCATATTTGGCTTA---GCCCTT 1009
 Db 146 AspLeuAlaLysLysValLysLysIlePheAsnLeuAspSerValArgLeuValSerTy 165
 OY 1010 GGGGTGGGAGAGAACTGAGTCAAGTCAAGCAAGTGGGCTGTGGTGGTGGTGGG 1069
 Db 166 GlyIuAsnAsnProLeuIleSerArg-----ValAlaIleCysGlyGlySerGly 182
 OY 1070 AGCAGCGTCTGAGAGGTGT-----GAGCTGACCTTTCACCTCAAGCTGAGATG 1120
 Db 183 GlnSerPheTyGlnIuAlaLeuThrLysGlyAlaGlnValIyrIleThrGlyAspIle 202
 OY 1121 TCCCATCATGATATCTTGGATGCTGCTCCCAAGAAATTAATGATCTCTGTGAACAC 1180
 Db 203 TyTyHisThrAlaGlnIuMetLeuThrAsnGlyLeuLeuAlaLeuAspPheGlyHis 222
 OY 1181 AGCAACACTGAAGAGAGCTTCTCTTCTGACCTCGAGATATGCTGGAT-----TTCAC 1234
 Db 223 ---HisIleGluValLeuPheValArgGlyLeuAlaGlnIuLysPheGlnThrTyPserCys 241
 OY 1235 TTGGAGAAAT---AAGATTAATATTATCTATCAGAGACTGACAGGAGCGCT 1282

Db 242 GlnGluAsnTrpAspIleThrIleLeuGlnSerGlnValAsnThrAsnPro 258
 RESULT 14
 ID P95777 PRELIMINARY; PRT; 262 AA.
 AC P95777;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE ORF4 protein.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=XC;
 RA MEDLINE=97175539; PubMed=9023194;
 RA Tsukloka Y., Yamashita Y., Oho T., Nakano Y., Koga T.;
 RT "Biological function of the dtdp-rhamnose synthesis pathway in
 RT Streptococcus mutans.";
 RL J. Bacteriol. 179:1126-1134(1997).
 DR EMBL; D78182; BAA11246.1;
 DR InterPro; IPR002678; DUF34.
 DR Pfam; PF01784; DUF34; 1.
 DR TIGRDB; TIGR00486; TIGR00486; 1.
 SQ SEQUENCE 262 AA; 29723 MW; 68E0D8BA6002DE3 CRC64;
 Alignment Scores:
 Pred. No.: 7.23e-08 Length: 262
 Score: 185.50 Matches: 65
 Percent Similarity: 36.28% Conservative: 50
 Best Local Similarity: 20.50% Mismatches: 101
 Query Match: 6.65% Indels: 101
 DB: 2 Gaps: 10
 US-09-745-506-74 (1-1553) x P95777 (1-262)
 OY 353 GTAATAACACTCTTCCTGACACATGACCTGAGAGAGTGTGAGAGAGTGTGCA 412
 Db 36 lIeLysArgLeuMetIleAlaLeuAspValArgIuThrThValAlaIleGlu 55
 OY 413 AAGAAGGACAGCACTTCTCTCTACACATCCGCTATCTCCGACCACTGAAGCGCAT 472
 Db 56 LysLysValAspLeuLeuIleValLysHisAlaProIlePheArgProLeuLysAsnLeu 75
 OY 473 ACCTGGAACACATGGAAGAGGCGCTGATCCGGCTGGAGAACAGAGTGGTATC 532
 Db 76 Val---GluThrAlaGlnAsnHisIleTyPheAsnLeuIleLysHisAlaVal 94
 OY 533 TACTCTCTCATACAGCCTATGATGCTGGCCCGCCAGGCGCTCAACACAGTGTGGCTAA 592
 Db 95 TyValSerHisThrAsnIleAspIleValProAspGlyLeuAsnSprTrpPheCysAsp 114
 OY 593 GGGCTTGGAGCTTGTACCTCCAGGCCCATACATCTCTCCAAAGCTCCCAACTACCTTACA 652
 Db 115 LeuLeuAspIleLysAsnArgIleLeuSerProSerLysAsp----- 129
 OY 653 GAGGAAACACAGAGTGAATTCACAGTAACTACACCAAGACCTGGACAAAGTCATG 712
 Db 129 ----- 129
 OY 713 TCTGACGTAAAGAAATGACGGTGTCTGTCACTTTCTTCTGTAGAGACTGTAA 772
 Db 129 ----- 129
 OY 773 GAGGAACAAACAGGATTATCTGATTTGACTCAGAAAGCTTGTGATGAGTGTAGAT 832
 Db 130 ----- 130
 OY 833 TTTCTTCCCGGAACAACACTTATCAGAAAGCAAAATTCGTCACTGAGAGAGCGCT 892

```

Db      131 Tyr-----
QY      893 TTGCTTACATGAAAGGAGCGTTATGACACATGAAATCTGTCCTGCGCA 952
Db      132 -----GlyIleGlyArgValGlyAspIle-----SerProLeuSerHeu 145
QY      953 ACCATGATGATTCGAATAAAGACACCTAAACCTAATCTCATTCGCTTA---GCCCTT 1009
Db      146 AspleuAlaLysValLysValLysIlePheAsnLeuAspSerValArgLeuValSerTyr 165
QY      1010 GGGGGGGGAGAACCTTTCAGCTTCACATCAAGTCAGCGCCCTGCGCTGCGCTCGG 1069
Db      166 GlyIleAsnAspProLeuIleSerArg-----IleAlaIleCysGlyLeuSerGly 182
QY      1070 AGCAGCGCTTCGACGGGTGT-----GAGCGTGCACCTTTCACCTACAGGTGAGATG 1120
Db      183 GlnSerPheTyrGlnGlnAlaLeuThrLysGlyAlaGlnValTyrIleThrGlyAspIle 202
QY      1121 TCCCATCTCATCTTTCGATGCTGCTGCTCCCAAGGATTAATGTCATCCTCTGAGAAC 1180
Db      203 TyrTyrHisThrAlaGlnGlnMetLeuThrAsnGlyLeuAlaLeuAspProGlyHis 222
QY      1181 AGCAACATGACAGCGCTTCTTCGACCTTCGAGATTCGATGCTGAT-----CTCAGC 1234
Db      223 ---HisIleGlnValLeuPheValArgLysLeuAlaGlnLysPheGlnThrTyrSerGys 241
QY      1235 TTGGAGAAAT---AAGATTAATATTTATCTCATCAGAGCTGACAGGACCTT 1282
Db      242 GlnGlnAsnTyrPaspIleThrIleLeuGlnSerGlnValAsnThrAsnPro 258

```

RESULT 15

```

Q9LAG2 PRELIMINARY; PRT; 265 AA.
AC
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 29.8 kDa protein.
OS Lactobacillus helveticus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1587;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=53/7;
RX MEDLINE=20120540; PubMed=10653753;
RA Savijoki K., Palva A.;
RT "Purification and molecular characterization of a tripeptidase from
   RT Lactobacillus helveticus";
RL Appl. Environ. Microbiol. 66:794-800(2000).
DR EMBL, AJ243321; CAB72937.1; -.
DR InterPro: IPR002678; D0F34.
DR Pfam: PF01784; D0F34, 1.
DR TIGRFAMs: TIGR00486; TIGR00486, 1.
KW Hypothetical protein.
SQ SEQUENCE 265 AA; 29850 MW; 716C9765F4860D17 CRC64;

```

Alignment Scores:

```

Pred. No.: 1.67e-07 Length: 265
Score: 181.50 Matches: 76
Percent Similarity: 37.32% Conservative: 55
Best Local Similarity: 21.65% Mismatches: 123
Query Match: 6.51% Indels: 97
DB: 2 Gaps: 12

```

US-09-745-506-74 (1-1553) x Q9LAG2 (1-265)

```

QY      251 TTGAAGGCTTCCTTCCTTCGATGATCTTCCTTCCTTCCTTCGAGAGCTGG 310
Db      4 ValIysAspIleValAsnArgLeuArgLysPheProGlnAspIleAlaSerLysGly 23
QY      311 GACATATGTTGATTACTGTGGAGAACCAAGCCACCATACATACATGTAATACACTCTTCG 370

```

```

Db      24 AspProValGlyMetGlnIle---GlySerMetGlnAlaAspValThrLysValMetThr 42
QY      371 ACCAATGACCTGATGAGAGATGATGAGAGGCTGCAAAAGAGACGACCTCAT 430
Db      43 ThrLeuAspValArgProGlnValValGlnAlaValGlnLysGlyValAsnPhelIle 62
QY      431 CTCCTCCATCACTCCGCTATCTTCGACCCATGAAAGCCCATPACCTGGAACATGAG 490
Db      63 ValSerHisHisProValMetPheArgProAlaArgAsnLeuAspPheAlaAspAlaGln 82
QY      491 GAGCCCGCTGGATGATCCGGGCTTCGGAACAGAGTCCGATCTCATCTCCATPACACC 550
Db      83 AsnAlaMetTyrGlyAsnIleIleLysAsnGlyIleThrValTyrSerIleHisThrAsn 102
QY      551 TATGATGCTGGCCCGCCAGGCGTCAACACTGCTGTTGCTAAAGGCTTGAGCTTGTACC 610
Db      103 SerAspLysAlaGlnAspGlySerLysAspIleProGlnAlaGlnLysIleuSerLysAsp 122
QY      611 TCCAGCCCATPACATCTTCCAAAGCTCCCAACTACCTACAGAGGAAACCCAGAGTA 670
Db      123 ValGlnProPhe-----
QY      671 GAATTCACGCTTAACATACACCAAGACCTGGACAAAGTCATGTCGACGTAAAGAAAT 730
Db      127 -----CysLeuAspAsp-----
QY      731 GAGCGTGTTCGTCACCTCTTTCTGCTGAGACTGTAAATGAGAAACAAACAGGAT 790
Db      131 AspGlyIleAlaMet-----GlyArgLysGlyArgLeuProGlnThr----- 144
QY      791 AATCTGAATTGACTACAGAGGCTTTGATGCGAGGTGTGATTTCTTCCCGGACAA 850
Db      145 -----MetThrAlaTyrAspPhe-----
QY      851 CAACCTTATCAGAAAGCGAAATTCGTACAGGAGAGACCTTCTCTCATATCTGGA 910
Db      151 ---AlaTyrTyrValLysGlnAsnMetGlyIleLys-----
QY      911 ATGGAGCGGTATGACACACTGATGATCTGTCTCCCTGGACACCATGATGATGCAATA 970
Db      162 MetaIaArgLeuIleThrAlaAspAsnAsnLysLysIleSerThr----- 176
QY      971 AAAACACACCTAAACCTATCTCATTTTCGCTTAGCCCTTGGGGTGGGAGAACCTTAG 1030
Db      176 -----
QY      1031 TCTCAAGTCAAAAGTCGTGGCCCTGTGCTGCTGTTCTGGAGACAGGCTTCTGACGGTGT 1090
Db      177 -----ValGlyPheIleCysGlyAspGlyLysTyrTyrThrArgAlaLeu 192
QY      1091 GAG-----GCTGACCTTTTACCTCACAGGTGAGATGTCCTCATCATGATCTTGGATGCT 1144
Db      193 AspAspHisLeuAspAlaPheIleThrGlyAspValTyrTyrHisValGlyHisAspMet 212
QY      1145 GCTTCCCAAGAAATGAATGATCTCTGTGAACACAGACCAACATGAGAGGCGCTTCT 1204
Db      213 IleSerSerGlyLeuThrValValAspProGlyHis---TyrThrGlnLysLeuPheLys 231
QY      1205 TCTGACCTTCGAGATATCTGATGATCT---CACTTGAGAGAAAT-----AAGTAATAATAT 1255
Db      232 TyrLysValIleTyrAspArgLeuLysLysTyrPasnGlnGlnAsnAsnTyrPasnValGlyVal 251
QY      1256 ATCTCATCAGAGACTGACAGGAGCCCTTTCAG 1288
Db      252 GlnLeuSerGlnValSerThrAsnProPheGln 262

```

Search completed: August 22, 2003, 14:05:51
Job time : 174 secs